Defense Technical Information Center
Compilation Part Notice

ADP014019

TITLE: Alternative Speech Sensors for Military Applications

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: Multi-modal Speech Recognition Workshop 2002

To order the complete compilation report, use: ADA415344

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP014015 thru ADP014027

Alternative Speech Sensors for Military Applications

U.S. Army Research Laboratory

Pete Fisher

6/10/02

Agenda

- Overview of current sensor technologies
- Possible future technologies
- Possible sensor fusion methods
- Military requirements
- Conclusion

Possible Methods for Improved ASR in Noise

- Reduce or eliminate noise through the processing of the speech signal based on properties of the signal components
- Detect speech without detecting noise
  - Many alternative sensors have reduced signal information
- Combinations of the above
  - Specialized sensors and processing
  - Multiple speech sensors and fusion of signals

Current Sensor Technologies

- Airborne acoustic microphones
- Contact acoustic microphones
- Bone conduction microphones
- Other alternative speech sensors

Airborne Acoustic Microphones

- Handheld microphones (Shure, etc)
- Headsets (Knowles, Shure, Telex, etc)
  - Noise canceling. close talking
- Super-directional microphones (Telex, etc)
  - Narrow band through beam forming
  - Linear arrays in a reinforcing pattern

Contact Acoustic Microphones

- Throat microphones (TEA, Genesys, Temco)
  - ARL Physiological Microphone
    - 32dB noise rejection
    - Acoustic response differs from a regular microphone
- Ear microphones (Jabra, Temco)
  - Some ear microphones are bone conduction
    - See next slide

Bone Conduction Microphones

- Navy bone conduction microphone
- Ear mounted bone conduction microphone
  - Invisio (TEA)
- Top of head bone conduction (Temco)
- Tooth mounted bone conduction microphone
  - Developed through a SBIR at CECOM

Other Alternative Speech Sensors

- Glottal Electromagnetic Micropower Sensor (GEMS)
  - Developed at Lawrence Livermore Nat. Labs
  - Commercial developer Aliph
  - Uses RADAR to measure internal motion
  - Reduced bandwidth
- Lip reading system (camera/computer)
  - Provides limited information, not a speech signal
  - Robust to noise

Other Alternative Speech Sensors

- Ultrasonic lip reader
  - Uses ultrasonic sensor to measure mouth opening
  - Have not been able to locate one of these devices, but have heard of them

Possible Future Technologies

- "Camera like" sensor that detects surface skin differently than tissues in the mouth
  - Would simplify detection of voiced speech
  - 3-5 and 8-12 micron FLIRs not suitable
  - Possibly some Near-Infrared technology?
- Novel vibration sensors
  - Technology?
    - Accelerometer? RADAR?

What to Sense?

- Vibration
  - Direct reading of speech or components
  - Close connection to avoid noise
- Alternatives?
  - Measure motion of speech articulators?
    - Tongue, teeth, glottis, sinuses
  - Modern jewelry?
    - Nose ring, cheek stud (microphones)
- Other methods?

Sensor Fusion

- Combining the outputs of one or more sensors to produce an improved speech signal
- Most appropriate in noisy environments where one or more sensors can be used to attempt to capture components of the speech signal while rejecting noise

Possible Sensor Fusion Methods

- Combine signals from multiple sensors in a cooperative fashion
  - Some non-standard speech sensors capture speech data while minimizing noise. but do not detect the full bandwidth of the speech signal
  - Could extract the cleanest spectral components of each sensor for input to ASR software

Possible Sensor Fusion Methods

- Use "clean speech" from noise robust sensors to remove noise from a primary sensor (airborne microphone)
  - Difference in secondary sensor signals and primary sensor signal is the noise (in the acoustic bands covered by the secondary sensors)
  - Could use correlation to remove noise that extends beyond the signal range of the sensor

Alternative Concept

- Work to improve a non-standard speech sensor and a matched ASR system to provide an integrated speech-in-noise package
  - Need a sensor with good noise rejection and "sufficient" signal capture capability
  - Need to tune the ASR engine to the peculiarities of the alternative speech sensor

Military Requirements

- Different for each application
  - Just like in the commercial world
- Selection of domain can be used to limit the problem
  - Command and control (C2) domain
    - Vocabulary of 1-5K words
    - Typically command phrases
    - Limited perplexity

24

Military Requirements (II)

- Most military environments will be noisy
  - Vehicles, people, weapons, generators, aircraft ...
- Capability to use existing microphones desirable in some cases
  - Communications via radio and vehicle intercoms
  - Difficulty of replacing all field equipment with improved or multi-modal speech sensors
  - Difficulty of getting more sensors on a soldier

What Do Military Users Want?

- They want a system that:
  - Works perfectly in all conditions
  - Weighs nothing
  - Is unbreakable
  - Does not interfere with their mission
  - Produces more energy than it uses
- Field soldiers are already overloaded
  - Make systems small (hand held), or make the software portable to platforms that are already carried by the soldier

Military Domains For SR

- C2 (command and control)
  - Constrained vocabulary, limited perplexity
  - "Tongue operated keyboard"
  - Electronic map navigation, radio settings
- Form completion
  - Repetitive task, limited vocabulary
  - Field reports, logistics (ordering supplies/ammo)
    - Might be performed over a low bandwidth field radio

Military Domains For SR (II)

- Information gathering
  - Vocabulary may not be constrained
    - User may have the option to enter free text fields with observations or other comments
  - Vehicle inspection, quality control
    - An actual military application of SR technology
- Monitoring of enemy communications
  - A much larger and more difficult application
  - Not amenable to application of alternative sensors

Conclusion

- There are a wide variety of alternative speech sensors available for exploitation for SR in military applications
- While many of these sensors do not detect the full range of human speech, their intrinsic noise rejection makes them useful
- Combinations of these alternative sensors may provide good solutions for the application of speech recognition in military environments

26